Some tests of the anticipatory mediated generalization model of acquired sample equivalence in pigeons' many-to-one matching

PETER J. URCUIOLI and KAREN M. LIONELLO-DeNOLF
Purdue University, West Lafayette, Indiana

Transfer-of-control tests typically show the development of acquired equivalence between samples occasioning the same comparison choice in pigeons' many-to-one matching-to-sample. Specifically, when some of those samples are later explicitly trained to occasion new comparison choices, the remaining samples immediately exert control over the new choices as well. In the present experiments, we examined whether or not this transfer effect depends on the order in which the various sample-comparison relations in training are learned. One group of pigeons initially acquired 0-delay many-to-one matching with four samples and two comparisons, followed by 0-delay matching with two of those samples and two new comparisons. Another group of pigeons learned the two-sample matching task first, followed by many-to-one matching. When subsequently tested for their ability to match the remaining samples from many-to-one matching to the comparisons used in the two-sample task, both groups showed comparable levels of transfer. These findings challenge the view that common anticipatory processes ostensibly arising from the samples in many-to-one matching are necessary mediators for the subsequent transfer effects indicative of acquired sample equivalence.

Categorization and acquired equivalence are among the more visible and active topics of research in animal cognition (Bovet & Vauclair, 1998; Hall, 1996; Schusterman, Reichmuth, & Kastak, 2000; Thompson & Oden, 2000; Wasserman & Astley, 1994; Zentall, 1998). A considerable amount of research dating back to the seminal studies of Herrnstein and Loveland (1964) has shown, for instance, that pigeons and other animals differentiate between classes of stimuli on the basis of perceptual similarity (Bhatt, Wasserman, Reynolds, & Knauss, 1988; Wasserman, Kiedinger, & Bhatt, 1988). Moreover, this learned categorical discrimination often generalizes immediately and accurately to novel instances of the trained categories. Thus, perceptually based categorization does not require language or other uniquely human capacities.

More interesting, perhaps, is categorization based on similarity in function or associative history. For instance, humans readily classify disparate-looking objects such as cars, boats, bicycles, airplanes, and trains together because each represents a form of transportation. Moreover, they exclude from this particular class any object that does not

serve a similar function, such as bananas, houses, trees, and books. This ostensibly more sophisticated form of categorization (Lea, 1984) also does not appear to require uniquely human capabilities, because it, too, is demonstrable in animals such as pigeons and rats (Honey & Hall, 1989; Johns & Williams, 1998; Urcuioli, Zentall, Jackson-Smith, & Steirn, 1989; Wasserman, DeVolder, & Coppage, 1992). One way to show such associatively based categorization in nonhuman animals is to initially train them to make different responses to, or to provide different reinforcers for, different sets of disparate stimuli (see, e.g., Astley & Wasserman, 1999; Urcuioli et al., 1989). After this initial training, new differential responses or reinforcers are then explicitly conditioned to some of the stimuli from each set. Finally, the remaining stimuli are tested to see whether they will also produce the new responses or conditioned reactions, despite the absence of any explicit conditioning to do so (cf. Shipley, 1935). The typical result has been that the remaining stimuli *do* have those effects, thus demonstrating that the stimuli comprising each of the original associative sets had become functionally equivalent (Goldiamond, 1962).

This general technique has been used successfully, for instance, to show acquired sample equivalence in pigeons' many-to-one matching-to-sample (Urcuioli et al., 1989; Wasserman et al., 1992). In this task, different-looking sample stimuli (e.g., a red hue and a white-on-black vertical line) initially occasion the same comparison choice response, whereas other disparate samples (e.g., a green hue and a white-on-black horizontal line) likewise occasion a common, but different, choice. The common choices rep-

This research was supported by National Institute of Mental Health Grant MH 56487 to P.J.U. The authors thank Lisa Huff, Michelle McIntosh, Leslie Peck, Jada Pierce, Jessica Rightsell, John Schmeling, and Steve Zuska for their assistance in running the experiments. Correspondence concerning this article should be addressed to P. J. Urcuioli, Department of Psychological Sciences, Purdue University, 1364 Psychology Building, West Lafayette, IN 47907-1364 (e-mail: uche@psych.purdue.edu).

resent the shared function or associative history for the members of each pair of samples. The acquired equivalence that might arise from this history can be assessed by whether or not new comparison choices learned to some of the samples immediately generalize to the remaining samples. Thus, following many-to-one training, one sample from each associative class (e.g., the red and green samples) is established as a conditional cue for a new comparison choice, after which the ability of the remaining samples (e.g., the vertical and horizontal lines) to occasion those new choices is tested. The typical finding has been that in testing, pigeons preferentially choose the same comparison that they have recently learned to make to the sample from the same associative class. In other words, if the red and vertical samples occasioned the same comparison choice in many-to-one matching, then a new choice later conditioned to the red sample preferentially occurs to the vertical sample in the transfer test, and likewise for the green and horizontal samples (Urcuioli, Zentall, & DeMarse, 1995; Urcuioli et al., 1989; Wasserman et al., 1992).

Although transfer effects such as these clearly establish associatively based categorization in pigeons, they do not, by themselves, address the mechanism(s) underlying the choice preferences observed in transfer. This issue was the primary focus of the present research.

One possible transfer mechanism was initially suggested by Hull (1939) in a paper devoted to stimulus equivalence. Applying his analysis to pigeons' many-to-one matching, Hull would argue that the samples associated with the same comparison choice eventually give rise to the (common) anticipation of that choice. When one sample from each associative class later occasions a new comparison choice during "reassignment training" (Wasserman et al., 1992), the originally conditioned anticipatory responses should, like the samples themselves, also acquire conditional stimulus control over the new choices. Consequently, the remaining samples should then be able to cue those choices in testing by virtue of the fact that they generate the same anticipatory responses that have just acquired control over the new choices during reassignment. In short, pigeons generalize new comparison choices from one set of samples (those appearing during reassignment training) to the remaining set (those appearing in testing) on the basis of common, anticipatory mediators. This account is known as anticipatory mediated generalization (see, e.g., Jenkins, 1963; Urcuioli, 1996).

The basic stimulus control features of this account are buttressed by the fact that pigeons routinely show transfer from one set of samples to another in other matching paradigms in which the potential mediator is made explicit—for example, by conditioning different response patterns to the sample stimuli (Urcuioli & DeMarse, 1994; Urcuioli & Honig, 1980; see also Peterson, 1984). The mediator hypothesized to explain acquired sample equivalence in many-to-one matching, however, is implicit. Nonetheless, it should be possible to independently evaluate the anticipatory mechanism by testing principled predictions derived from it. The present experiments provided such tests.

The rationale for the experimental designs was straightforward. If common anticipatory responses (viz., common prospective codes—Honig & Thompson, 1982) truly underlie transfer of new choices between samples with a common associative history, then those anticipatory responses must be "in place" prior to, or at least during, acquisition of those choices. That way, they, like the samples that purportedly generate them, can cue which new comparison choice to make. Consequently, if the new choices are learned to a subset of the samples *before* the many-to-one task itself, then transfer to the remaining samples should be either absent or greatly reduced. The reason is that the samples alone will have acquired control over those choices *before* any common anticipatory responses can become conditioned to them.

We tested this prediction in both experiments described below. The design of each experiment was virtually identical to that of the other, with procedural details differing only slightly, although in a manner that may be behaviorally and theoretically important. These details along with their potential implications will be highlighted at appropriate places in the discussion of each experiment.

EXPERIMENT 1

Table 1 summarizes the design of Experiment 1. Each letter in the notation beneath the various experimental phases represents a pair of samples or comparisons, with samples designated by the letter to the left of the dashes and comparisons by the letter to the right.

Group First initially learns many-to-one matching—that is, to match one set of comparisons (B) to two sets of samples (A and C). Subsequently, these pigeons learn to match one set of samples (A) to new comparisons (D) prior to testing their ability to match the remaining (C) samples to the D comparisons. Group Second is trained on exactly the same sample-comparison relations as is Group First, but in reverse order: A–D matching followed by many-to-one (A–B, C–B) matching. Group Control learns all of the same training relations (A–B, A–D, and C–B) as do the other two groups but without any *concurrent* training involving different samples matched to a common pair of comparisons. The latter group initially learns one-to-many matching (see, e.g., Urcuioli & Zentall, 1993), in which a single pair of samples (A) is matched to two different pairs of comparisons (B and D). Afterwards, it learns to match one pair of comparisons (B) to new samples (C) prior to testing on the C–D relations. Given that past research has typically found no evidence of transfer following one-to-many training in pigeons (Urcuioli & Zentall, 1993; Urcuioli et al., 1995), Group Control was included to assess the possibility of reduced-level transfer in Group Second vis-à-vis Group First.

Note that this experimental design ensures that all groups are trained on the same baseline relations (A–B, A–D, and C–B) and are tested on the same set of relations (C–D). Consequently, any between-group differences in testing must be due to the order and/or context in which the individual training relations are learned.

Table 1
The General Design of Experiment 1 (All Three Groups)
and Experiment 2 (Groups First and Second Only)

Group	Training		Testing
	Phase 1	Phase 2	
First	A–B	A–D	C–D
	C–B		
Second	A–D	A–B	C–D
	A–D	C–B	
Control	A–B	C–B	C–D
	A–D		

Note—Each letter represents a pair of samples or comparisons. Letters to the left of each dash are samples; those to the right are comparisons.

Again, the mediated generalization model outlined above predicts that transfer indicative of acquired sample equivalence should be apparent in Group First but not in Group Second (nor in Group Control; cf. Urcuioli et al., 1995).

Method

Subjects

A total of 34 experimentally naive White Carneaux pigeons obtained from the Palmetto Pigeon Plant (Sumter, SC) began the study. The pigeons were housed individually in stainless steel wire mesh cages in a colony room with a 14:10-h light:dark cycle with light on at 0700 h. Each bird was maintained at 80% of its free-feeding body weight throughout the experiment by restricting its feeding primarily to the experimental sessions. On days in which sessions were not run, each bird received a maintenance allotment of food in the home cage, where water and health grit were always available. Prior to the start of the experiment, the birds were randomly divided into six groups consisting of the three main groups shown in Table 1, each of which, in turn, was divided into consistent and inconsistent subgroups for testing (described below). Four of the six resulting groups were each assigned 6 subjects, and the remaining two were each assigned 5. One pigeon died during the course of the study.

Apparatus

Two standard conditioning chambers for pigeons, each consisting of a BRS/LVE Model PIP-016 three-key panel inside a BRS/LVE Model SEC-002 enclosure, were used. Each key measured 2.5 cm in diameter and was separated by 5.7 cm (center to center) from the adjacent key(s). The three keys were positioned in a horizontal row 7.5 cm from the top of the panel, and each was equipped with an inline stimulus projector mounted directly behind it. The projectors could display red, green, and white homogeneous fields, a small white dot and a small inverted white triangle on a black background, and sets of three vertical and three horizontal white lines on black backgrounds (BRS/LVE Pattern No. 692). A food hopper containing mixed grain was accessible, when raised, through a 5.8-cm square opening located 9 cm below the center key. The raised hopper was lit by a miniature bulb mounted at the top of the back wall of the metal housing covering the hopper. A GE No. 1829 bulb located above the center key and partially covered by a metal housing, the opening of which directed light toward the ceiling, provided chamber illumination. A continuously running blower fan mounted on the outside of the chamber provided ventilation and masking noise. Two Gateway 2000 386SX/25 computers controlled and recorded all experimental events.

Procedure

Preliminary training. After the pigeons were initially trained to eat consistently and quickly from a periodically raised and lit food

hopper, the method of successive approximations was used to shape pecking to the center response key when lit by the white inverted triangle. Over the next five 60-trial sessions, the pigeon learned to peck at each of the stimuli that would later serve as samples or as comparisons in matching-to-sample. The first two sessions consisted of center-key presentations of the dot and white stimuli and of red and green hues, respectively. The next two sessions consisted of center-key presentations of red and green in conjunction with side-key presentations of red and green hues or vertical and horizontal lines, respectively. In the final session, the dot and white appeared on the center key, and red and green hues appeared on each side key. In all sessions, each scheduled stimulus appeared an equal number of times at its respective location. A single peck at the lit key on a trial immediately turned off the stimulus and produced a 3-sec access to mixed grain. Successive stimulus presentations were separated by a 10-sec intertrial interval (ITI). The houselight was turned off during the first 9 sec of the ITI, and it was turned back on for the last 1 sec of the ITI and for the duration of the trial (i.e., until the end of the reinforcement cycle).

The next 10 preliminary training sessions involved only center-key presentations of red and green (5 sessions), or dot and white (subsequent 5 sessions) with the required number of pecks to produce reinforcement gradually raised from 2 to 10 over each 5-session block.

Matching-to-sample training. Table 2 depicts the same training and testing contingencies shown in Table 1, but in expanded form. S1–S4 represent the different sample stimuli (red, green, dot, and white, respectively), and C1–C4 represent the different comparison alternatives (red, green, vertical lines, and horizontal lines, respectively). The two hue comparisons (C1 and C2) always appeared together as choice alternatives as did the two line comparisons (C3 and C4). The “+” indicates the correct choice alternative for each trial type.

In *Phase 1*, Group First was trained on many-to-one matching, in which pecks to the red comparison (C1) were reinforced on trials beginning with either the red (S1) or the dot (S3) sample and pecks to the green comparison (C2) were reinforced on trials beginning with either the green (S2) or the white (S4) sample. Group Second was trained on two-sample/two-alternative matching with red and green samples (S1 and S2) and vertical and horizontal line comparisons (C3 and C4). For half of the Group Second birds, a vertical comparison choice was reinforced on red-sample trials and a horizontal

Table 2
An Expanded Schematic of Experiment 1 (All Three Groups)
and Experiment 2 (Groups First And Second Only)

Group	Training		Testing
	Phase 1	Phase 2	
First	S1 → C1+	S1 → C3+	S3 → C3+
	S2 → C2+		
	S3 → C1+		
	S4 → C2+		
Second	S1 → C3+	S1 → C1+	S3 → C3+
	S2 → C4+	S2 → C2+	
		S3 → C1+	
		S4 → C2+	
Control	S1 → C1+	S3 → C1+	S3 → C3+
	S2 → C2+		
	S1 → C3+		
	S2 → C4+		

Note—S1–S4 are sample stimuli, C1–C4 are comparison stimuli, “+” indicates the correct comparison choice following each sample. Incorrect comparison choices are not shown. Testing contingencies shown are for the consistent subgroups; sample–comparison relations were reversed for the inconsistent subgroups (not shown).

comparison choice was reinforced on green-sample trials; for the remaining birds, these contingencies were reversed. Finally, Group Control learned to match red and green samples (S1 and S2) to both red and green comparisons (C1 and C2) and to vertical- and horizontal-line comparisons (C3 and C4) in Phase 1. For all of the birds in this group, red comparison choices were reinforced on red-sample trials and green comparison choices were reinforced on green-sample trials. However, the reinforced line comparison choices vis-à-vis the red and green samples were partially counterbalanced across the 10 birds in this group.

Each Phase 1 session consisted of 96 trials for Groups First and Control and 100 trials for Group Second, with each sample appearing the same number of times in a session. For Group Control, both sets of comparison alternatives (hues and lines) also appeared equally often in each session. Likewise, the left-right side-key configurations of the single set of comparisons (Groups First and Second) or the two sets (Group Control) appeared equally often following each sample. The various types of matching trials were presented in pseudo-random order in each session, with the constraint that no combination of sample and comparisons appear more than three times in a row.

Each matching trial began with presentation of one of the sample stimuli on the center key. Ten pecks to the sample turned it off and immediately produced the two comparison alternatives on the adjacent side keys (i.e., a 0-delay procedure). A single peck to either comparison then turned off both comparisons and yielded food reinforcement if the correct comparison was pecked, or a time-out period (houselight turned off) of equal duration if the incorrect comparison was pecked. Reinforcement duration was constant within a session for each bird, but varied from 2 to 6 sec across birds and sessions in such a way as to maintain each bird's body weight as close to the 80% value as possible. A 10-sec ITI, the first 9 sec of which were spent in darkness, immediately followed reinforcement or time-out. The houselight was then turned on for the last 1 sec of the ITI and remained on until the end of the next trial.

Phase 1 training for each bird was run for a minimum of 10 sessions and until the bird reached an overall accuracy level of 90% correct or better for 5 of 6 consecutive sessions. In addition, choice accuracy for each set of samples for the birds in Group First, or for each set of comparisons for the birds in Group Control, had to be at least 87.5% correct over the same sessions. Once a Group First or Group Control bird met the accuracy criteria, it received an additional 20 overtraining sessions on its respective Phase 1 task.

In Phase 2, Group First acquired the same two-sample/two-alternative matching task that was used in Phase 1 training for Group Second. Similarly, Group Second learned the many-to-one task acquired by Group First in Phase 1. For Group Control, Phase 2 training consisted of daily 100-trial sessions in which the dot and white (S3 and S4) samples were matched to red and green comparisons (C1 and C2), respectively. All remaining details of the Phase 2 tasks (including the performance criteria) were identical to those described for Phase 1 training.

For all groups, each Phase 2 training session alternated on a daily basis with continued training on the Phase 1 matching task in order to maintain accurate performances on the originally learned task (cf. Urcuioli et al., 1995; Urcuioli et al., 1989, Experiment 2). In other words, each training session on two-sample/two-alternative matching for Groups First and Control alternated with continued training on its previously acquired many-to-one and one-to-many matching task, respectively. Similarly, each Phase 2 many-to-one training session for Group Second alternated with continued training on its previously acquired two-sample matching task. For all birds, Phase 2 training continued for a minimum of 10 sessions and until they met the aforementioned accuracy criteria. At that point, birds in Group First and Group Control were moved to the test phase of the experiment. Birds in Group Second, however, were given an additional 20 overtraining sessions on many-to-one matching (with continued daily alternation with its Phase 1 matching task) in order to equate its total many-to-one training with that given to Group First.

Matching-to-sample testing. In testing, birds in all three groups matched the dot and white samples (S3 and S4) to vertical- and horizontal-line comparisons (C3 and C4). Each test session consisted of 100 matching trials divided equally among the four possible trial types (two samples $\times$ two left-right configurations of the comparisons). Trials were randomized, with the sole constraint that no one trial type could appear more than three times in a row. Reinforcement continued to be provided for each correct choice response, and all other details of these test sessions were identical to those described for the two-sample/two-alternative matching task used in training. Each bird was tested for a minimum of 10 sessions and until it reached a criterion of 90% or greater accuracy for 5 of 6 consecutive test sessions.

Half of the birds in each group were tested with matching contingencies that were consistent with any acquired equivalence that may have developed in training between the samples associated with the same correct comparison choice. The other half of the birds in each group were tested with matching contingencies that were inconsistent with any acquired sample equivalence that may have previously developed. For example, for birds tested in the consistent subgroups, if the red and dot samples (S1 and S3) had both cued the same comparison choice (C1) in training, and if the red (S1) sample had also been established as a conditional cue for choosing the vertical-line comparison (C3), then choosing the vertical-line comparison (C3) was reinforced after the dot sample (S3) in testing. Likewise, for the green and white samples (S2 and S4), the horizontal-line choice (C4) that was reinforced in training after green (S2) was also reinforced after the white (S4) in testing. These consistent contingencies are depicted in Table 2 in the "Testing" column.

By contrast, for birds assigned to the inconsistent subgroups, the test contingencies were the opposite of those described above (not shown in Table 2). For example, if the red and dot samples (S1 and S3) had both cued the same comparison choice (C1) in training, and if the red (S1) sample had been established as a conditional cue for choosing the vertical-line comparison (C3), then choosing the *horizontal*-line comparison (C4) was reinforced after the dot sample (S3) in testing. Similarly, the *vertical*-line choice (C3) was reinforced after the white sample (S4) in testing.

Predictions. If an acquired equivalence had developed in training between the samples occasioning the same reinforced comparison choice, then the consistent subgroups should match at higher levels of accuracy during testing than would the inconsistent subgroups. Indeed, first-session accuracy for consistent birds should be above the level expected by chance (50%), whereas it should be below that level for the inconsistent birds. On the other hand, if training in any group had not produced an acquired sample equivalence, then performances should be comparable across the respective (consistent and inconsistent) subgroups. It is important to note that counterbalancing of the line-comparison choices with the red and green samples during training ensured that the matching contingencies represented in the consistent and inconsistent subgroups were equated within each group. Consequently, any differences between consistent and inconsistent test performances within a group cannot be attributed to the specific sample-comparison relations reinforced in testing.

Statistical analyses. Analysis of variance (ANOVA) was initially used to evaluate differences within and/or between groups. When the ANOVA indicated a significant overall effect, post hoc contrasts across the group means were evaluated with the use of the methods, tabled *F* values, and inferential procedures described by Rodger (1975a, 1975b). Rodger's tabled *F* values ensure that the expected proportion of all null contrasts rejected in error will not exceed .05, the Type I error rate for all analyses reported here.

Results

Training Performance

The rates at which the groups acquired accurate baseline matching performances depended on the nature of

Table 3
Mean Percentages of Correct Choice Responses
Over the Last Five Sessions With Each Set of
Training Relations for Each Group in Experiment 1

Group	Training Relations		
	A–B	C–B	A–D
First–Consistent	97.3	97.9	94.6
First–Inconsistent	98.8	96.3	95.3
Second–Consistent	98.4	98.4	97.3
Second–Inconsistent	98.0	95.6	96.2
Control–Consistent	97.8	95.2	94.1
Control–Inconsistent	95.0	93.0	93.0

Note—Group data are divided according to whether the subsequent transfer test involved sample–comparison relations consistent or inconsistent with the training relations.

the task (i.e., many-to-one, one-to-many, or two-sample/two-alternative matching) and on whether it was learned in Phase 1 or Phase 2. For example, acquisition to criterion of one-to-many matching by Group Control took longer on average (34.6 sessions) than acquisition of many-to-one matching by Groups First and Second (12.1 sessions). For the latter birds, those acquiring many-to-one matching in Phase 2 (Group Second) reached criterion sooner on average (10.6 sessions) than those acquiring it in Phase 1 (Group First: 13.6 sessions).

Despite the variation in acquisition rates, terminal levels of matching accuracy on the three sets of baseline re-

lations (A–B, C–B, and A–D) were uniformly very high. Baseline accuracies averaged over the last five training sessions with each set of relations ranged from 93.6% to 98.2% across groups. For A–B matching, these accuracies were 98.0%, 98.2%, and 96.4% for Groups First, Second, and Control, respectively [$F(2,30) = 1.90$]. For C–B matching, mean accuracy was slightly but significantly lower in Group Control (94.1%) than in Groups First and Second (each at 97.1%) [$F(2,30) = 3.55$]. For A–D matching, Group Second (96.8%) was significantly more accurate than Group Control (93.6%) [$F(2,30) = 4.05$], with Group First performing at an intermediate level (95.0%) [$F(2,30) = 0.02$].

The between-group differences in terminal C–B and A–D accuracies were of little concern, however, because the differences were small and because acquired equivalence was assessed by *within*-group comparisons of transfer performances as a function of whether the tested (C–D) relations were consistent or inconsistent with the training relations. More critical, then, are the terminal accuracy data across subgroups, and these data are summarized in Table 3. None of the within-group differences in performance on any of the three baseline relations was statistically significant (all F s < 2.76).

Transfer Performance

The left panel of Figure 1 shows matching accuracy on the first transfer test session averaged across birds in

Figure 1. Average percentage of correct choice responses by each group on the first transfer test session in Experiment 1 (left panel) and Experiment 2 (right panel). Data are shown separately for the consistent and inconsistent subgroups.

each group and test condition. Figures 2 and 3 show the corresponding individual first-session test performances.

For both Groups First and Second, average matching accuracy on the first transfer session was substantially higher for birds tested in the consistent than in the inconsistent test condition. Specifically, the Group First birds tested with sample-comparison relations consistent with an acquired sample equivalence chose the correct comparison on 67.3% of all first-session test trials, whereas those tested with inconsistent relations chose correctly on only 45.3% of those trials. The corresponding figures for the

Group Second birds were 68.2% and 47.2%, respectively. Birds initially trained on one-to-many matching (Group Control) showed the opposite pattern: those tested on consistent relations matched correctly on 49.8% of the first-session test trials, whereas those tested on inconsistent relations matched correctly on 61.2% of those trials.

The individual-subject data show that for Group First, there was very little overlap between first-session performances of birds tested on consistent versus inconsistent relations. All 6 birds in this group that were tested with consistent relations matched at accuracy levels above

Figure 2. Percentages of correct choice responses on the first transfer test session of Experiment 1 for individual subjects in Group First (top panel) and Group Second (bottom panel). Data for birds tested with matching relations consistent versus inconsistent with acquired sample equivalence are shown in the left and right panels, respectively.

Figure 3. Percentages of correct choice responses on the first transfer test session of Experiment 1 for individual subjects in Group Control. Data for birds tested with matching relations consistent versus inconsistent with acquired sample equivalence are shown in the left and right panels, respectively.

50% (range: 52%–86%), whereas 5 of the 6 birds that were tested with inconsistent relations matched at levels near or below 50% (range: 32%–53%). For Group Second, there was more overlap across conditions. Specifically, although all of the Group Second birds in the consistent test condition matched at levels greater than 50% (range: 52%–91%), only 2 birds in the inconsistent test conditions were below chance (range: 31%–63%). For Group Control, the range of accuracies for birds tested in the consistent condition was 41%–60% versus 48%–77% in the inconsistent condition.

A one-way ANOVA on the first-session means showed a significant overall effect across the six treatment conditions (3 groups $\times$ 2 test conditions) [$F(5,27) = 4.08$]. Pairwise comparisons using the error mean square from the overall ANOVA were then used to assess within-group differences in performance as function of consistent versus inconsistent testing. A significant effect of test condition was found in Group First [$F(5,27) = 1.96$] and in Group Second [$F(5,27) = 1.62$], but not in Group Control [$F(5,27) = 0.44$]. The former two statistical findings are of specific interest to the theoretical issue motivating this experiment, and indicate that transfer indicative of acquired sample equivalence was present in both groups receiving many-to-one training.

An additional set of contrasts compared first-session matching accuracies across groups as well as across test conditions. These contrasts showed that the level of accuracy in the consistent test condition for Groups First and Second did not differ significantly from one another [$F(5,27) = 0.00$], nor from the inconsistent condition for Group Control [$F(5,27) = 0.21$]. Similarly, first-session matching accuracies did not differ between the inconsis-

tent conditions for Groups First and Second [$F(5,27) = 0.01$], and this level was comparable to that observed in the consistent condition for Group Control [$F(5,27) = 0.06$]. The overall level of accuracy in the former three conditions did, however, significantly differ from that in the latter three [$F(5,27) = 3.62$]. These statistical decisions imply the following ordering of accuracies on the first test session: Group First–Consistent = Group Second–Consistent = Group Control–Inconsistent > Group First–Inconsistent = Group Second–Inconsistent = Group Control–Consistent (Rodger, 1975b, Equation 23).

Figure 4 plots performances on the transfer (C–D) relations over 10 repeated test sessions. The data for Groups First and Second, both of which received training on many-to-one matching, are shown in the top panel, and the corresponding data for Group Control, whose training involved one-to-many matching, are shown in the bottom panel. The filled symbols plot the data from the consistent test conditions; open symbols plot the data from the inconsistent conditions.

Three features of the repeated-session transfer performances depicted in Figure 4 stand out. One is that the initial within-group differences in matching accuracy between the consistent and inconsistent conditions in Groups First and Second remained throughout most of the 10 test sessions. Another is that within a particular test condition (consistent or inconsistent), matching accuracy on the C–D relations was very comparable in Groups First and Second. Finally, the within-group differences seen in Groups First and Second were reversed in Group Control.

The test data from the birds trained on many-to-one matching were further analyzed in a split-plot ANOVA using the test condition as part of the group identifier (i.e.,

First-Consistent, First-Inconsistent, Second-Consistent, and Second-Inconsistent). This ANOVA revealed significant effects of group [$F(3,19) = 5.29$] and of test session [$F(9,171) = 66.74$], and a significant group $\times$ session interaction [$F(27,171) = 3.10$]. Post hoc contrasts on performances averaged across the 10 test sessions showed that matching accuracy in Group First-Consistent (87.2%) did not differ from that in Group Second-Consistent (83.0%) [$F(3,19) = 0.20$]. Likewise, there was no significant difference in the corresponding accu-

racies of Group First-Inconsistent (70.7%) and Group Second-Inconsistent (69.2%), [$F(3,19) = 0.02$]. The two consistent conditions together, however, yielded significantly higher levels of accuracy averaged over the 10 test sessions than did the two inconsistent conditions [$F(3,19) = 5.03$].

A similar ANOVA on the test data from the birds that had been trained on one-to-many matching showed significant effects of test condition [$F(1,9) = 10.90$] and test session [$F(9,72) = 37.46$], but no interaction [$F(9,72) = 0.57$].

Figure 4. Average percentage of correct choice responses over the first 10 transfer test sessions in Experiment 1 for pigeons trained on many-to-one matching (Groups First and Second, top panel) and on one-to-many matching (Group Control, bottom panel). Filled symbols plot the data from the consistent test conditions; open symbols plot the data from the inconsistent test conditions.

Matching accuracy averaged over the 10 test sessions was 69.6% in Group Control–Consistent and 82.6% in Group Control–Inconsistent.

Discussion

The transfer to novel sample–comparison relations following many-to-one training seen in this experiment is not a novel result (see, e.g., Urcuioli et al., 1995; Urcuioli et al., 1989; Wasserman et al., 1992). However, the finding that the consistent versus inconsistent difference in transfer materializes even when many-to-one training *follows* training in which different contingencies are applied to one member of each (future) sample class *is* novel. Group Second, in other words, showed transfer effects indicative of acquired sample equivalence very much like those observed in Group First, which learned its many-to-one task prior to reassignment training. In short, the order in which the two sets of training relations were learned (viz., those constituting many-to-one matching and those in the two-sample matching task) does not matter. With both training orders, different comparison choices that were explicitly conditioned to only two samples from many-to-one matching occurred preferentially in testing to the remaining, associatively related samples. Theoretically, this is the most interesting finding of this experiment, and it will be discussed more fully momentarily.

Although the order in which training on many-to-one matching and the two-sample “reassignment” task appears to be inconsequential, acquired sample equivalence does appear to depend on *concurrent* training of the common comparison associations. In other words, when those associations are learned to the same high degree of accuracy but in separate training sessions, as exemplified by the one-to-many sequence of Group Control (cf.

Table 2), no evidence of acquired sample equivalence (i.e., transfer of new choices across samples) appears.

Indeed, transfer following one-to-many training in Group Control was significantly more accurate when the contingencies in testing were *inconsistent* rather than consistent with acquired sample equivalence, a rather unexpected result. We believe, however, that this difference probably reflects random factors rather than a true effect, given that other published findings have failed to reveal a consistent versus inconsistent transfer difference following one-to-many matching (Urcuioli & Zentall, 1993; Urcuioli et al., 1995). Moreover, recent unpublished data from our lab also show no test condition differences following both standard one-to-many training and a variation of it involving additional, different combinations of the choice alternatives (cf. Saunders & Green, 1999). Thus, even though other assays have suggested stimulus class formation in pigeons’ one-to-many matching (e.g., Bonardi, Rey, Richmond, & Hall, 1993; Zentall et al., 1992), data obtained from two-choice training and testing procedures have yet to produce clear, reproducible evidence for equivalence.

In contrast, differences in transfer after many-to-one versus one-to-many training like those obtained here (see also Urcuioli et al., 1995) are consistently reproducible in other paradigms and with other species. For example, Spradlin and Saunders (1986) and Saunders, Drake, and Spradlin (1999) reported stronger and/or more consistent transfer effects in children and adults with mental retardation after many-to-one (“comparison-as-node”) training than after one-to-many (“sample-as-node”) training, using tests in which subjects were required to match either the sample stimuli or the comparison stimuli from training to one another. Similarly, Delamater and Joseph (2000,

Table 4
Prospective (Anticipatory) and Retrospective Mediated Generalization

Group	Phase 1 Training	Phase 2 Training + Theoretics	Testing + Theoretics
Prospective Mediation			
First	A — B
	C — B	A — D	C — D
Second	A — D
		A — B	C — D
		C — B	
Retrospective Mediation			
First	A — B
	C — B	A — D	C — D
Second	A — D
		A — B	C — D
		C — B

Experiment 1) found stronger equivalence class formation in college students trained on many-to-one symbolic matching than on one-to-many matching, using total versus partial reversals of the previously learned sample-response relations as an assay. Finally, Johns and Williams (1998) found better evidence of equivalence class formation between different stimuli signaling a common unconditional stimulus (a many-to-one relation) than between different stimuli each of which followed a common antecedent unconditional stimulus (a one-to-many relation), using rats trained and tested in Pavlovian preparations.

Differences in transfer following many-to-one versus one-to-many training in pigeons' matching-to-sample can be derived, with a few additional assumptions, from a model of anticipatory mediated generalization (Urcuioli, 1996). However, the data from Group Second challenge this model by questioning the idea that transfer of new choices across the many-to-one samples requires a covert conditional cue reflecting the common anticipation elicited by those samples. The top half of Table 4 shows this problem more clearly by using the task notation from Table 1 plus the hypothesized anticipatory links purported to develop during Phase 2 training and to be present during testing. The lowercase letters represent the anticipation of the comparison choices conditioned to one or both sets of samples. When many-to-one matching is learned initially (Group First), the common anticipation of "b" to the A and C samples provides common (across-phase) "b"-D links that should support transfer from the samples participating in Phase 2 reassignment (A) to those appearing in testing (C). But those common mediating links ought to be absent in Group Second. Instead, when many-to-one matching is learned *after* the A-D task, anticipatory processes conditioned to the A samples in Phase 1 should generate "d"-B links in Phase 2. But those links provide no mechanism for the C samples to cue the D choices in testing, because the "b" anticipation that should eventually develop to the C samples in Phase 2 (not shown) is not linked to the previously learned D comparisons.

However, one might reasonably argue that since Phase 2 training alternated on a daily basis with continued training on the Phase 1 task, this afforded the opportunity for "b"-D links to develop in Group Second too. The required assumptions are that (1) a "b" anticipation developed to both A and C samples as the many-to-one task was learned in this group (reasonable), and (2) once the "b" anticipation began to occur regularly to the A samples, the alternating training on the Phase 1 A-D task allowed the "b"-D links to form (questionable). We list the latter assumption as questionable because the postulated "b"-D links in Group Second would represent redundant conditional stimulus control, given that the A samples had already acquired strong control over the D choices during Phase 1 training. Thus, even granting the first assumption and the carryover of the "b" anticipation to the alternating A-D task, the previously learned A-D relations should block the development of "b"-D links (see, e.g., vom Saal & Jenkins, 1970). In other words, the already es-

tablished conditional stimulus control by the A samples should prevent or greatly diminish acquisition of redundant conditional stimulus control by the anticipatory processes to which those samples later give rise. Comparison of the transfer results from Group Second with those of Group First does not support this prediction.

Nonetheless, rejecting anticipatory mediated generalization as a necessary ingredient for transfer to novel sample-comparison relations following many-to-one training would be more compelling if we could experimentally make the possibility of "b"-D links even less likely. The next experiment was designed with this goal in mind.

EXPERIMENT 2

In Experiment 2, we replicated the two many-to-one conditions of Experiment 1 (cf. Tables 1 and 2) with one notable difference. After each pigeon acquired its Phase 1 matching task, training during Phase 2 was *not* alternated with continued training on the Phase 1 contingencies. In other words, as Group First learned to match two samples from many-to-one matching to two new comparison choices during Phase 2 reassignment, there were no intervening refresher sessions on many-to-one matching. Likewise, as Group Second acquired and received overtraining on many-to-one matching during Phase 2, the two-sample/two-choice matching task that they had learned in Phase 1 was discontinued.

This relatively small procedural change has potentially important theoretical consequences. Specifically, eliminating the alternation between the two matching tasks allows one to avoid, or at least to greatly reduce (see the Procedure section), the possibility that any sample-specific anticipatory responses conditioned during many-to-one matching in Group Second could acquire control over its *previously learned* choices. Anticipatory ("b") responses conditioned to the many-to-one samples in Group First, on the other hand, might have some limited opportunity to acquire control over new (D) comparison choices because this group learns the many-to-one matching task first. However, even in Group First, the development of anticipatory ("b"-D) control depends critically on two variables: (1) the speed with which the new (D) choices are learned in Phase 2, and (2) the continued occurrence of the anticipatory ("b") responses to the A samples even though the A-B matching component of the Phase 1 (many-to-one) task is completely discontinued during Phase 2 reassignment. Although the former can be directly measured, the latter, unfortunately, cannot.

Nonetheless, under the most favorable circumstances in which residual anticipatory responses in Group First are maintained long enough to acquire some control over the D choices in Phase 2 training, these birds should show some evidence of transfer of those new choices to the remaining samples from many-to-one matching. Under less favorable circumstances (e.g., the D choices are acquired after the "b" anticipations no longer occur), no transfer would be expected in this group. Similarly, no transfer

should be apparent in Group Second, because any “b” anticipations arising from their many-to-one samples during the Phase 2 training cannot cue the D comparisons that these birds have already learned.

Method

Subjects and Apparatus

Sixteen experimentally naive White Carneaux pigeons obtained from the Palmetto Pigeon Plant participated in Experiment 2. The birds were randomly divided into four groups consisting of the many-to-one groups shown in Tables 1 and 2 (Groups First and Second), split again into consistent and inconsistent subgroups. All birds were housed and maintained in the same fashion as in Experiment 1. The apparatus and control equipment were also the same as those used previously.

Procedure

Preliminary training. Each bird went through exactly the same preliminary training sequence with the same stimuli as did the birds in Experiment 1.

Matching-to-sample training. *Phase 1* matching-to-sample training was identical to that described for Groups First and Second in Experiment 1. In short, each bird in Group First was initially trained on 0-delay many-to-one matching with red, green, dot, and white (S1–S4) samples and red and green comparison alternatives (C1 and C2) until it reached a criterion of five of six consecutive sessions with 90% or better overall matching accuracy and 87.5% or better accuracy with each pair of samples. The birds then received an additional 20 overtraining sessions on this task. Each bird in Group Second, on the other hand, learned to match red and green samples (S1 and S2) to vertical- and horizontal-line comparisons (C3 and C4) during Phase 1 training. These birds were trained for a minimum of 10 sessions and until they matched at accuracy levels of 90% or greater for 5 of 6 consecutive sessions. All other details of Phase 1 training, including counterbalancing of the sample–comparison contingencies in Group Second, were identical to those described in Experiment 1.

During *Phase 2* training, Group First learned the two sample/two-alternative matching task that Group Second had acquired in Phase 1, whereas Group Second learned the many-to-one matching task initially acquired by Group First. In this experiment, however, each Phase 2 training session was run successively across days; there was no daily alternation, in other words, between the Phase 2 and Phase 1 matching tasks.

After each bird completed Phase 2, it received *refresher* training on its Phase 1 task in preparation for testing. Although the inclusion of refresher sessions would appear to compromise the stated purpose of Experiment 2, they were absolutely necessary in order to verify that each bird could continue to match accurately with the samples (Group First) or the comparisons (Group Second) that subsequently appeared during the transfer test. (Stated otherwise, if the baseline performances of any bird prior to testing were inaccurate, its test results would be uninterpretable.) Many-to-one matching constituted refresher training for Group First because the two-sample/two-alternative matching task intervened between its Phase 1 training task and testing. By contrast, two-sample/two-alternative matching constituted refresher training for Group Second because many-to-one acquisition and overtraining intervened between its Phase 1 training and testing. In order to equate both the nature of the training relations and performance levels across groups immediately prior to the first test session, the following refresher sequences and performance criteria were used: Group First—one session at criterion levels of accuracy on many-to-one matching, one session at criterion on two-sample/two-alternative matching, and one session at criterion on many-to-one matching; Group Second—one session at criterion levels of accuracy on two-sample/two-alternative matching and one session at criterion on many-to-one matching.

Matching-to-sample testing. During transfer testing, all birds matched the dot and white samples (S3 and S4) to vertical- and horizontal-line comparisons (C3 and C4). One half of the birds in each group were tested with matching contingencies consistent with any acquired sample equivalence that may have developed during many-to-one training, whereas the remaining half were tested with contingencies inconsistent with acquired equivalence. All details regarding these contingencies and the testing procedure itself were identical to those in Experiment 1.

Results

Training Performances

Although Groups First and Second differed in the rates at which they acquired each training task, these differences appeared mostly to reflect the ordinal position in which a particular matching task was learned. Specifically, both many-to-one matching and the two-sample matching task were learned more rapidly when acquired during Phase 2 than during Phase 1. For instance, the average number of sessions to criterion levels of accuracy on many-to-one matching was 19.25 for Group First during Phase 1; it was 12.5 for Group Second during Phase 2. The corresponding figures for two-sample/two-alternative matching acquisition were 10.75 for Group First during Phase 2 and 16.12 for Group Second during Phase 1.

More importantly, terminal levels of accuracy on the three sets of baseline matching relations (A–B, C–B, and A–D) constituting the two training tasks were uniformly high in both groups. Baseline accuracies averaged over the last five training sessions with each set of relations ranged from 95.4% to 99.1% correct in Group First and from 93.0% to 98.4% correct in Group Second. There were no significant between-group differences on any baseline matching relation [all $F_s(1, 14) < 3.48$].

Table 5 shows the mean percentages of correct choice responses for each group on each set of training relations from the final refresher session involving that relation. Each group's data have been split according to whether birds were later tested in the consistent or in the inconsistent transfer condition. Once again, baseline accuracies were very high on all training relations. The only significant between-group difference involved the C–B training relations: Group Second matched more accurately than Group First [$F(1, 14) = 17.97$]. However, accuracies *within* each group (i.e., across consistent and inconsistent sub-

Table 5
Mean Percentages of Correct Choice Responses on Each Set of Training Relations During the Final Refresher Sessions for Each Group in Experiment 2

Group	Training Relations		
	A–B	C–B	A–D
First–Consistent	95.3	96.4	95.2
First–Inconsistent	96.4	96.4	94.2
Second–Consistent	97.4	99.0	94.8
Second–Inconsistent	96.9	100.0	96.0

Note—Group data are divided according to whether the subsequent transfer test involved sample–comparison relations consistent or inconsistent with the training relations.

groups) were comparable on all sets of training relations [all $F_s(1,6) < 3.02$].

The average number of refresher sessions needed for birds to recover Phase 1 performances to criterion levels of accuracy was 1.5 for Group First (range: 1–2 sessions) on many-to-one matching and 3.9 for Group Second (range: 3–5 sessions) on two-sample/two-alternative matching. This between-group difference probably reflects the fact that Phase 2 training lasted longer on the average in Group Second (viz., 37 sessions, which included the 20 overtraining sessions) than in Group First (15 sessions).

Transfer Performance

The right panel of Figure 1 shows the average accuracy for each group on the first test session. Figure 5 shows the corresponding individual performances. For both groups, mean first-session accuracy was higher in the consistent test condition than in the inconsistent test condition. For Group First, every bird tested in the consistent condition matched above the level expected by chance (range: 53%–84% correct), although the above chance level of accuracy for 2 birds was marginal. By contrast, matching accuracy for 3 of the 4 Group First birds tested in the in-

Figure 5. Percentages of correct choice responses on the first transfer test session of Experiment 2 for individual subjects in Group First (top panel) and Group Second (bottom panel). Data for birds tested with matching relations consistent versus inconsistent with acquired sample equivalence are shown in the left and right panels, respectively.

consistent condition were well below chance (42%, 33%, and 43% correct); the remaining bird was 51% correct. For Group Second, 3 of the 4 birds tested in the consistent condition matched at accuracy levels above chance on their first test session (65%, 63%, and 65% correct); the remaining bird was 49% correct. First-session accuracies for two Group Second–Inconsistent birds, on the other hand, were well below chance (37% and 41% correct); the remaining 2 birds performed at 49% and 51%, respectively. In passing, it is perhaps noteworthy that the Group Second–Consistent bird that performed the poorest (49% correct) was also the least accurate of all 16 birds in the experiment on its final refresher session.

An ANOVA on the first-session test results showed a significant overall between-group/condition difference [$F(3,12) = 5.61$]. Post hoc contrasts across the group means showed that mean first-session accuracy did not differ between Groups First–Consistent (64.2%) and Second–Consistent (60.5%) [$F(3,12) = .11$], nor did it differ between Groups First–Inconsistent (42.2%) and Second–Inconsistent (44.5%) [$F(3,12) = .04$]. However, overall first-session accuracy for the two consistent test conditions was significantly greater than accuracy for the two inconsistent test conditions [$F(3,12) = 5.46$]. These statistical decisions imply the following ordering of first-session transfer performances: Group First–Consistent = Group Second–Consistent > Group First–Inconsistent = Group Second–Inconsistent (Rodger, 1975b, Equation 23). In short, the consistent versus inconsistent difference in transfer to novel (C–D) sample–comparison relations was evident in both groups and comparable across groups.

Figure 6 plots the performances by each group on the transfer (C–D) relations over the first 10 test sessions. The filled symbols plot the data from the consistent test conditions; open symbols plot the data from the inconsistent test conditions. These data resemble those from Experiment 1 (cf. Figure 4, top panel) in that the consistent versus inconsistent differences in performances observed within each group were maintained over 4–5 sessions. In addition, matching accuracy on the C–D relations within each test condition (consistent and inconsistent) was reasonably comparable across groups. The most noticeable difference between these repeated-test-session data and those of Experiment 1 was that acquisition of the C–D test relations in this experiment progressed relatively rapidly, especially in both of the inconsistent test conditions.

The data plotted in Figure 6 were initially analyzed in a split-plot ANOVA, using test condition as part of the group identifier (i.e., First–Consistent, First–Inconsistent, Second–Consistent, and Second–Inconsistent). This ANOVA showed significant effects of group [$F(3,12) = 4.97$] and test session [$F(9,108) = 87.93$], and a significant group $\times$ test session interaction [$F(27,108) = 2.33$]. Post hoc contrasts across the group means averaged over all 10 sessions showed that Groups First–Consistent and Second–Consistent did not differ from one another [$F(3,12) = 0.36$], nor did Groups First–Inconsistent and Second–Inconsistent [$F(3,12) = 0.00$]. However, matching by the two consistent groups was significantly more accurate than matching by the two inconsistent groups [$F(3,12) = 4.63$].

Separate split-plot ANOVAs comparing just the two consistent groups and just the two inconsistent groups revealed

Figure 6. Average percentage of correct choice responses over the first 10 transfer test sessions in Experiment 2 for Groups First and Second. Filled symbols plot the data from the consistent test conditions; open symbols plot the data from the inconsistent test conditions.

no significant between-group differences [$F_s(1,6) = 1.93$ and 0.01 , respectively], and no significant group $\times$ test session interactions [$F_s(9,54) = 0.94$ and 1.14 , respectively].

Discussion

The findings of Experiment 2 provide a stronger challenge to the anticipatory mediated generalization account of acquired equivalence than do those of Experiment 1. The reason is that the absence of alternating training between many-to-one matching and the two-sample (reassignment) task substantially decreases the chances that the choices in the latter task (D) will come under the control of the anticipatory ("b") processes presumably conditioned to the A and C samples in the former task. But did this procedural change *preclude* the development of "b"–D links?

There may be good reason to claim that Experiment 2 did not completely eliminate such links in Group Second, given that these birds required roughly four refresher (albeit nonalternating) sessions to recover accurate performance on two-sample (A–D) matching following completion of many-to-one training. Moreover, average matching accuracy on this group's first refresher session was only 64.9% correct (range: 54%–81%). Thus, the demonstrable loss of conditional stimulus control by the A samples caused by the lengthy Phase 2 many-to-one training in this group may have permitted the opportunity for any "b" anticipation conditioned to the A samples during Phase 2 to acquire conditional stimulus control as similar control was being reestablished to the A samples themselves. Unfortunately, as this group's first refresher session results show, it is imperative to verify that accuracy on the baseline training relations is high prior to testing. Without such verification (i.e., without refresher training to recover baseline accuracies), the likely result of chance-level test accuracy might only reflect the loss of sample or comparison control *per se* (cf. Carter & Eckerman, 1975).

Although some interpretive ambiguity surrounds the effects of Group Second's refresher training, we can also judge the viability of anticipatory mediated generalization by closer examination of the training and testing results from Group First. Recall that for this group, the postulated "b"–D links necessary for transfer must develop relatively early in the transition from Phase 1 many-to-one matching to Phase 2 matching. We say "early" because a reasonable expectation is that anticipation of the B comparisons from many-to-one matching should cease rather quickly during Phase 2 training when the A samples are no longer followed by them but only by the D comparisons instead. This contrasts with the situation in Experiment 1 where every Phase 2 training session alternated with the many-to-one task in which the hypothesized "b" anticipations were established and maintained.

But acquisition of accurate A–D matching by Group First was not particularly rapid during Phase 2 training. For example, its average accuracy on the *fifth* Phase 2 training session was only 69% correct, with 4 of the 8 birds still matching close to chance (viz., between 50% and 57% correct). Furthermore, two of the Group First birds that were most accurate on Session 5 (87% and 90%) showed the

poorest transfer in their respective test conditions (cf. the first individual-subject bar in the Group First–Consistent and the Group First–Inconsistent data in Figure 5). On the other hand, the Group First–Consistent bird showing the strongest first-session transfer performance (84% correct) was still matching at chance levels (51%) on its fifth Phase 2 training session. Clearly, there was no obvious correlation between transfer performances and the speed with which the A–D training relations were learned in this group.

GENERAL DISCUSSION

The present experiments provide a number of salient results and implications regarding acquired sample equivalence in pigeons' matching-to-sample. First, the consistent versus inconsistent difference in transfer to novel sample–comparison relations following training on 0-delay many-to-one matching once again demonstrates that samples associated with the same correct choice become functionally equivalent. The formation of the acquired equivalence classes was seen by the fact that after such training, the samples were interchangeable with one another in new discriminations. In other words, after one sample from each presumed class developed control over a new (or different) comparison choice via explicit (reinforced) training in two-sample/two-alternative delayed matching, the other sample in the class also exerted the same control despite the absence of any prior, explicitly reinforced training to do so (see also Spradlin, Cotter, & Baxley, 1973). This finding fits classic definitions of a stimulus class of the sort offered by Goldiamond (1962): "Once a class is established, contingencies applied to one member of a class will tend to affect other members of the class" (p. 303).

Second, the common comparison associations that establish acquired sample equivalence must be trained concurrently. When pigeons learn to match different samples to the same comparison but learn those relations in separate training sessions and different contexts, those samples are not interchangeable with one another in the control of new/different choices. This was demonstrated by the one-to-many (Group Control) results of Experiment 1.

Third, the interchangeability or substitutability of the samples from many-to-one matching in their control of new/different choices does not depend on when conditional stimulus control over those choices by a subset of the samples is established. The results of both Experiment 1 and Experiment 2 show that those choices can be explicitly trained either after or before many-to-one training. In either case, new conditional relations between the "untrained" samples and the new/different choices emerge, as is shown by the consistent–inconsistent differences in transfer in both Group First and Group Second.

Before further theoretical discussion, we should note that although identity matching (matching red and green hues to red and green comparisons) was part of the training relations in each experiment, it would be a mistake to strictly interpret those relations as A–A and as integral

to obtaining transfer of control. Regarding the latter point, Urcuioli et al. (1995, Experiment 1), Urcuioli, DeMarse, and Zentall (1994), and Wasserman et al. (1992) have all reported transfer effects of the sort reported here following many-to-one training in which all training relations were symbolic. Regarding the former point, some of our other published data (Lionello & Urcuioli, 1998) show very clearly that the same stimulus on different keys is functionally a different stimulus for the pigeon. In other words, given that the samples and comparisons appear in different locations, the sample "A" and the comparison "A" in the A–A relation are probably more accurately represented by the symbolic or nonidentity A–B description.

We have indicated that appealing to common anticipatory processes to explain transfer-of-control across the samples from many-to-one matching is theoretically questionable. Granted, nothing in our results definitively rules out such anticipatory mediated generalization. But this account is clearly stretched by the findings that transfer occurs (1) independently of when many-to-one training is conducted, (2) without session-to-session overlap between training of the many-to-one matching relations and those involving the new choices, and (3) despite the fact that acquisition of the new sample–comparison relations from which the transfer relations emerge is slow enough that postulating continued anticipatory processes is implausible.

An alternative mediational account of acquired equivalence that has been suggested by Wasserman and DeVolder (1993) and by Zentall (1998) is that samples associated with the same comparison act as retrieval cues for each other. This retrospective account, in which "a" and "c" represent the retrieved representations of the A and C samples, respectively, appears in the bottom half of Table 4. In order to predict transfer of performance to the C–D test relations, this account must assume that the remembered version of a sample not currently present is similar to the actual sample itself, an assumption supported by other data (see, e.g., Holland & Forbes, 1982). The actual A–D relations learned by Group First during Phase 2, then, should permit the retrieved representation of A ("a") by the C samples in testing to cue the D choices (i.e., "a"–D is very much like A–D). Likewise, for Group Second, the retrieval of "a" by the C samples in testing should cue the D choices, given that the A samples themselves were established as conditional cues for D choices during Phase 1. The attractive feature of this account is that the hypothesized mediational processes ostensibly depend only on many-to-one training per se and not when such training is conducted.

Although retrospective mediation can account for the present results, it still leaves the question of the mechanism by which the A and C samples activate the representations of their associatively related counterparts unanswered. At a minimum, that mechanism would appear to necessitate both symmetrical (or bi-directional) and transitive associations. For example, a C sample could

retrieve the corresponding A sample with which it shares a common (B) comparison if many-to-one (A–B, C–B) training yielded training relations that were symmetrical (e.g., B–A). Thus, the explicitly trained C–B relation combined with the symmetrical B–A relation should yield C–A if the two relations in question were also transitive. Unfortunately, these assumptions are immediately challenged by the general lack of evidence in pigeons for symmetry and transitivity in two-choice conditional discrimination tasks.

Maybe the precise nature of the mediated generalization processes producing these observed transfer effects will, in time, be specified by more discriminating empirical tests. Then again, given past problems with mediated generalization accounts of equivalence, perhaps the belief that such tests will eventually materialize is overly optimistic. Moreover, it is worth keeping in mind that the number of assumptions and steps necessary to derive the predicted data could easily make any resulting model(s) cumbersome and/or implausible (Jenkins, 1963; Sidman, 1994). With these caveats, Sidman's position that equivalence is best viewed as a basic stimulus function—not derivable, in other words, from other processes—looks increasingly attractive.

REFERENCES

- ASTLEY, S., & WASSERMAN, E. A. (1999). Superordinate category formation in pigeons: Associations with a common delay or probability of food reinforcement makes perceptually dissimilar stimuli functionally equivalent. *Journal of Experimental Psychology: Animal Behavior Processes*, **18**, 193–207.
- BHATT, R. S., WASSERMAN, E. A., REYNOLDS, W. F., JR., & KNAUSS, K. S. (1988). Conceptual behavior in pigeons: Categorization of both familiar and novel examples from four classes of natural and artificial stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, **14**, 219–234.
- BONARDI, C., REY, V., RICHMOND, M., & HALL, G. (1993). Acquired equivalence of cues in pigeon autoshaping: Effects of training with common consequences and with common antecedents. *Animal Learning & Behavior*, **21**, 369–376.
- BOVET, D., & VAUCLAIR, J. (1998). Functional categorization of objects and of their pictures in baboons (*Papio anubis*). *Learning & Motivation*, **29**, 309–322.
- CARTER, D. E., & ECKERMAN, D. A. (1975). Symbolic matching by pigeons: Rate of learning complex discrimination predicted from simple discriminations. *Science*, **187**, 662–664.
- DELAMATER, A. R., & JOSEPH, P. (2000). Common coding in symbolic matching tasks with humans: Training with common consequence or antecedent. *Quarterly Journal of Experimental Psychology*, **53B**, 255–273.
- GOLDIAMOND, I. (1962). Perception. In A. J. Bachrach (Ed.), *Experimental foundations of clinical psychology* (pp. 280–340). New York: Basic Books.
- HALL, G. (1996). Learning about associatively activated stimulus representations: Implications for acquired equivalence and perceptual learning. *Animal Learning & Behavior*, **24**, 233–255.
- HERRNSTEIN, R. J., & LOVELAND, D. H. (1964). Complex visual concept in the pigeon. *Science*, **146**, 549–551.
- HOLLAND, P. C., & FORBES, D. T. (1982). Control of conditional discrimination performance by CS-evoked event representations. *Animal Learning & Behavior*, **10**, 249–256.
- HONEY, R. C., & HALL, G. (1989). The acquired equivalence and distinctiveness of cues. *Journal of Experimental Psychology: Animal Behavior Processes*, **15**, 338–346.

- HONIG, W. K., & THOMPSON, R. K. R. (1982). Retrospective and prospective processing in animal working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 16, pp. 239-284). New York: Academic Press.
- HULL, C. L. (1939). The problem of stimulus equivalence in behavior theory. *Psychological Review*, **46**, 9-30.
- JENKINS, J. J. (1963). Mediated associations: Paradigms and situations. In C. N. Cofer & B. S. Musgrave (Eds.), *Verbal behavior and learning: Problems and processes* (pp. 210-245). New York: McGraw-Hill.
- JOHNS, K. W., & WILLIAMS, D. A. (1998). Acquired equivalence learning with antecedent and consequent unconditioned stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, **24**, 3-14.
- LEA, S. E. G. (1984). In what sense do pigeons learn concepts? In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 263-276). Hillsdale, NJ: Erlbaum.
- LIONELLO, K. M., & URCUIOLI, P. J. (1998). Control by sample location in pigeons' matching to sample. *Journal of the Experimental Analysis of Behavior*, **70**, 235-251.
- PETERSON, G. B. (1984). How expectancies guide behavior. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 135-148). Hillsdale, NJ: Erlbaum.
- RODGER, R. S. (1975a). The number of non-zero, *post hoc* contrasts from ANOVA and error rate: I. *British Journal of Mathematical and Statistical Psychology*, **28**, 71-78.
- RODGER, R. S. (1975b). Setting rejection rate for contrasts selected *post hoc* when some nulls are false. *British Journal of Mathematical and Statistical Psychology*, **28**, 214-232.
- SAUNDERS, R. R., DRAKE, K. M., & SPRADLIN, J. E. (1999). Equivalence class establishment, expansion, and modification in preschool children. *Journal of the Experimental Analysis of Behavior*, **71**, 195-214.
- SAUNDERS, R. R., & GREEN, G. (1999). A discrimination analysis of training-structure effects on stimulus equivalence outcomes. *Journal of the Experimental Analysis of Behavior*, **72**, 117-137.
- SCHUSTERMAN, R. J., REICHMUTH, C. J., & KASTAK, D. (2000). How animals classify friends and foes. *Current Directions in Psychological Science*, **9**, 1-6.
- SHIPLEY, W. C. (1935). Indirect conditioning. *Journal of General Psychology*, **12**, 337-357.
- SIDMAN, M. (1994). *Equivalence relations and behavior: A research story*. Boston: Authors Cooperative.
- SPRADLIN, J. E., COTTER, V. W., & BAXLEY, N. (1973). Establishing a conditional discrimination without direct training: A study of transfer with retarded adolescents. *American Journal of Mental Deficiency*, **77**, 556-566.
- SPRADLIN, J. E., & SAUNDERS, R. R. (1986). The development of stimulus classes using match-to-sample procedures: Sample classification vs. comparison classification. *Analysis & Intervention in Developmental Disabilities*, **6**, 41-58.
- THOMPSON, R. K. R., & ODEN, D. L. (2000). Categorical perception and conceptual judgments by nonhuman primates: The paleological monkey and the analogical ape. *Cognitive Science*, **24**, 363-396.
- URCUIOLI, P. J. (1996). Acquired equivalences and mediated generalization in pigeon's matching-to-sample. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 55-70). Amsterdam: Elsevier.
- URCUIOLI, P. J., & DEMARSE, T. (1994). On the relationship between differential outcomes and differential sample responding in matching-to-sample. *Journal of Experimental Psychology: Animal Behavior Processes*, **20**, 249-263.
- URCUIOLI, P. J., DEMARSE, T., & ZENTALL, T. R. (1994). Some properties of many-to-one matching with hue, response, and food samples: Retention and mediated transfer. *Learning & Motivation*, **28**, 175-200.
- URCUIOLI, P. J., & HONIG, W. K. (1980). Control of choice in conditional discriminations by sample-specific behaviors. *Journal of Experimental Psychology: Animal Behavior Processes*, **6**, 251-277.
- URCUIOLI, P. J., & ZENTALL, T. R. (1993). A test of comparison-stimulus substitutability following one-to-many matching by pigeons. *Psychological Record*, **43**, 725-744.
- URCUIOLI, P. J., ZENTALL, T. R., & DEMARSE, T. (1995). Transfer to derived sample-comparison relations by pigeons following many-to-one versus one-to-many matching with identical training relations. *Quarterly Journal of Experimental Psychology*, **48B**, 158-178.
- URCUIOLI, P. J., ZENTALL, T. R., JACKSON-SMITH, P., & STEIRN, J. N. (1989). Evidence for common coding in many-to-one matching: Retention, intertrial interference, and transfer. *Journal of Experimental Psychology: Animal Behavior Processes*, **15**, 264-273.
- VOM SAAL, W., & JENKINS, H. M. (1970). Blocking the development of stimulus control. *Learning & Motivation*, **1**, 52-64.
- WASSERMAN, E. A., & ASTLEY, S. L. (1994). A behavioral analysis of concepts: Its application to pigeons and children. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 31, pp. 73-132). New York: Academic Press.
- WASSERMAN, E. A., & DEVOLDER, C. L. (1993). Similarity- and non-similarity-based conceptualization in children and pigeons. *Psychological Record*, **43**, 779-793.
- WASSERMAN, E. A., DEVOLDER, C. L., & COPPAGE, D. J. (1992). Non-similarity-based conceptualization in pigeons. *Psychological Science*, **3**, 374-379.
- WASSERMAN, E. A., KIEDINGER, R. E., & BHATT, R. S. (1988). Conceptual behavior in pigeons: Categories, subcategories, and pseudocategories. *Journal of Experimental Psychology: Animal Behavior Processes*, **14**, 235-246.
- ZENTALL, T. R. (1998). Symbolic representation in animals: Emergent stimulus relations in conditional discrimination learning. *Animal Learning & Behavior*, **26**, 363-377.
- ZENTALL, T. R., SHERBURNE, L. M., STEIRN, J. N., RANDALL, C. K., ROPER, K. L., & URCUIOLI, P. J. (1992). Common coding in pigeons: Partial versus total reversals of one-to-many conditional discriminations. *Animal Learning & Behavior*, **20**, 373-381.

(Manuscript received June 12, 2000;
revision accepted for publication May 8, 2001.)